Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 35 (2014) 1307 – 1314

Procedia Computer Science

18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems - KES2014

# Development of finger motion skill learning support system based on data gloves

Kazuma Iwasako, Masato Soga, Hirokazu Taki*

Faculty of System Engineering, Wakayama University, Wakayama 640-8510, Japan

**Abstract**

Recently, more and more hearing person starts learning sign language. For finger motion skill learning such as sign language, it is important for an expert to give objective advice to a learner, because the learner is not able to find any motion errors in his/her sign language. In other words the learner needs the self-education tool of sign language which gives good objective advice. However, there are few studies about such a system. In this paper, we developed a finger motion skill learning support system using data gloves. This system helps a learner to recognize motion errors intuitively by himself/herself by overlaying skilled person's 3D models with learner's 3D models of hands and fingers.



Keywords: Finger Motion; Skill Learning; Learning Support; Data Gloves; 3D Model;

## 1. Introduction

For finger motion skill leaning such as sign language, it is important for an expert to give objective advice to a leaner, because it is hard for the learner to find any motion errors in his/her sign language. In case of self-education, the learner may learn wrong motion without any advice. In other words the system needs the function to give objective advice to learners supporting their self-education of sign language.

Most of the studies in this field only supports recognizing or translating sign language for a learner. There are a few studies in which a motion learning support system for sign language was developed. Kuribayashi et al. (2004) has developed such a system using data gloves[1]. Kitagawa et al. (2013) has developed another system using Kinect and bending sensor[2]. Kuribayashis' system[1] sets Japanese fingerspelling as a target. The system induces a learner to master a skill of performing the correct hand motion. The system shows the

* Corresponding author. Tel.: +81-73-457-8457.
*E-mail address:* soga@sys.wakayama-u.ac.jp.


doi:10.1016/j.procs.2014.08.167

correct hand motion using text and CG animation. The system allows the learner to learn each character of Japanese fingerspelling one by one, however, series of characters such as words are unsupported. Kitagawas' system[2] covers 182 words of sign language. The system is using Kinect for getting potential relations with their hands and bodies. However, Kinect cannot get motion data of their hands in detail. That is why their system uses data gloves with bending sensors to get learners' hands shape data.

In this paper, we developed a finger motion skill learning support system using data gloves. This system allows a learner to recognize errors intuitively by himself/herself by overlaying skilled person's 3D models with learner's 3D models of hands and fingers. This system also provides feedback of errors to the learner, calculated by matching process with their hand form data.

In our laboratory, we developed the system which reproduces finger motion by 3D models using magnetic position sensors[3]. However, there are some problems such that it takes long time to put them on, and the learner feels burdensome while he/she uses them. It is caused by the fact that the system has 12 sensors on one hand and they are wired. That is why we adopt data gloves which are easy for a learner to put on and to take off.

In the following sections, section 2 gives composition of the developed system. Section 3 explains Japanese fingerspelling the system targets. Section 4 gives the way to render 3D models of hands and fingers. Section 5 explains diagnosis system that calculates errors. Section 6 shows results of questionnaire to evaluate the system. In section 7, we consider the system.

## 2. System composition

The system we developed consists of data gloves to get learner's finger motion and a PC to feedback (Fig. 1).

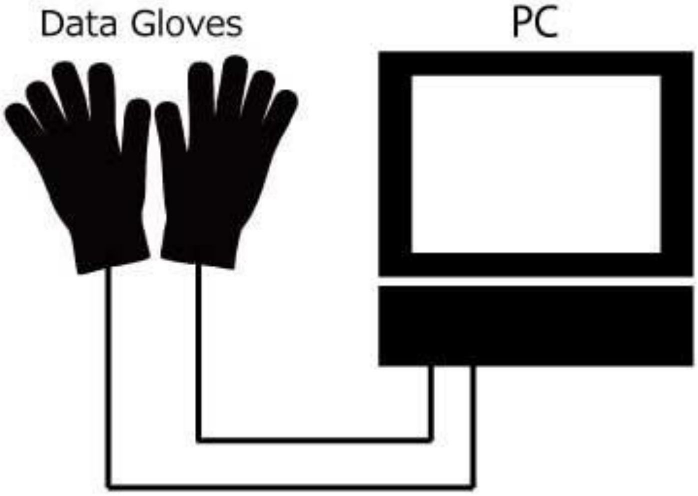


Fig. 1. The system's composition

### 2.1. 5DT data glove 14 ultra

It is a motion capture system for hands and fingers that 5DT Inc. developed. It has total 14 sensors on second joints and third joints of each finger, and the sensors return fractional value between 0 and 1 as bend states of each joint.

### *2.2. Development environment*

The system has been developed on a computer machine that Windows 7 is loaded. We use the development environment and language as shown below.

- Environment : Microsoft Visual Studio C++ 2010 Express
- Language : C++
- GUI : Windows Form Application
- Library : OpenGL, GLMetaseq

### *2.3. Window composition*

The system comprises three components, a 3D Model View, an Advice View and Control Panels (Fig. 2).

3D Model View shows skilled person's models and learner's models. Skilled person's model reproduces skilled motions as the target motion made by skilled person's motion data measured in advance. Learner's model reproduces learner's motion on a real-time basis made by motion data from data gloves that learner puts on. It can change a view point by dragging on the 3D Model View.

Control Panels allows the learner to select a fingerspelling he/she learns, change speed to render model, calculate errors of learner's motion and set learner's models to be shown or hidden.

Advice View has two text areas. One shows advice to improve learner's motion, and indicates some fingers which have large errors in comparison with skilled person's motion. Another shows score of learner's motion, and improvement value from last score.

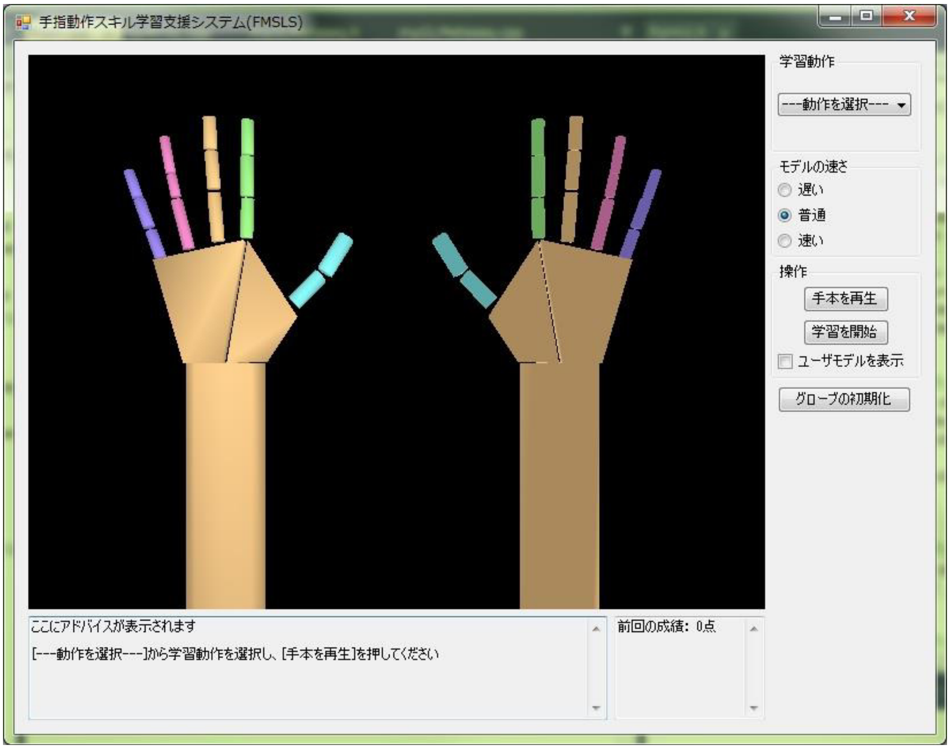


Fig. 2. The system's GUI

## 3. Targeted motion

### *3.1. About Japanese fingerspelling*

We have developed the system which targets the motion of Japanese fingerspelling. Japanese fingerspelling is a language that associates speaker's hand shapes with Japanese unvoiced consonants (Fig. 3). Generally, it is mainly used to express proper nouns like one's name, or words when the speaker does not know how to express it by sign language. In addition to their hand shapes, by moving their arms, Japanese fingerspelling allows us to express voiced consonants, p-sound, syllabic nasal in Japanese and assimilated sound. However, these expressions are exempt from our targets, because the data glove cannot get hand position data.


Fig. 3. Examples of fingerspelling

### *3.2. Reason of selection*

As stated above, the hand shapes are important for recognizing Japanese fingerspelling.

In addition, Japanese fingerspelling is useful for evaluating some of the functions of the system. Because there are several expressions of Japanese fingerspelling using similar hand shapes and bending fingers. For example, in Japanese fingerspelling, there are similar expressions between 'e' and 'sa' or between 'chi' and 'tsu'. Therefore, we can evaluate the system by confirming whether or not the learner is able to recognize such fingerspelling motions of 3D models exactly. We can also evaluate the system by confirming whether or not the system is able to recognize expression the learner performed. In addition, the expression 'mo' has a motion that bend thumb and index fingers. Therefore, we can check by the expression whether the system is able to diagnose the motions in which the learner has to change hand shapes to express the fingerspelling.

## 4. Render 3D models of hands and fingers

3D finger models we used are made of MQO files which were made by a 3D modeler software "Metasequoia". OpenGL and GLMetaseq were used to reproduce the models. GLMetaseq is a library to handle OpenGL.

### *4.1. The way for reproduction of motions*

To reproduce motions, we use the rotation angles which are calculated from data measured by a sensor. Values of angles are needed to rotate 3D models with OpenGL. The values of the data measured by sensors are not degrees but fractional values of joint flex, therefore, they need to be converted from fractional values to value of angles.

The method of converting data from fractional values by the sensors to value of angles is as follows.

Sensors are set 0.0 when fingers are extended, and set 1.0 when fingers are most flexed. They get fractional values which are in the ratio of a flexed angle to the most flexed angle. Thus a flexed angle *A(s)* is gotten by the following equation (1).

$$A(s) = s\{A(Smax) - A(Smin)\} + A(Smin) \quad (1)$$

where s is the value of a sensor, Smin is the minimum value of a sensor, and Smax is the maximum value of a sensor.

Using a flexed angle *A(s)* to achieve the joint rotation, 3D models is able to reproduce a motion of hands and fingers.

### *4.2. Modification of GLMetaseq*

To reproduce a motion of hands and fingers by 3D models, we have modified GLMetaseq (Table 1). The modification was caused by rotational process of OpenGL, which applies to each MQO_MODEL. Each 3D model we made consists of 14 objects. When the system reads data of 3D models, before it is modified, all objects in the data are stored into a structural value MQO_MODEL. In this case, a rotational process of OpenGL is applied to all objects, because an entire object of a 3D model is rotated, and the system cannot reproduce hand shapes. In another hand, after that, each object is stored into an element of structural array MQO_MODEL. It allows to apply a rotational process to an optional object, and allows to access to an element. Therefore, to apply a rotational process to the object which position is corresponded with the sensor, the system is able to reproduce the learner's hand shape.

Table 1. Modification of GLMetaseq

| | |
|---|---|
| Before | Store objects in a same file into a structural value MQO_MODEL |
| After | Store objects in a same file into each element of structural array MQO_MODEL[size]. (size = number of objects) |

## 5. The diagnosis system

### *5.1. The way to calculate learner's error*

Error is calculated on each finger, based on a similarity between skilled person's fingerspelling and learner's one by DP matching. Fingerspelling motion data is recorded in chronological order of angles, which is translated from values of sensors. The system calculates similarities by DP matching between skilled person and a learner. It is Error values that similarities are added up on each fingers, and normalized. If an error value was larger than threshold, the system advises the learner with text that some fingers have large error.

### *5.2. Scoring motions the learner performed*

A score for motions is a value for similarity of whole motion of hands and fingers. It is calculated by totalizing and normalizing similarities of all joints. The maximum score is 100 point, and the minimum score is 0 point.

## 6. Evaluation experiment

Subjects replied to a questionnaire after they used the system (Table 2). On the questionnaire, Q1 to Q13 are 5-point scale evaluation (1 is the lowest, 5 is the highest), Q14 is free description.

Table2. Questions

| No. | Question |
|---|---|
| Q1 | 3D models reproduce your hands and fingers exactly. |
| Q2 | The size of 3D models is suitable. |
| Q3 | When 3D models are overlaid, it is easy to check their motions. |
| Q4 | By overlaying 3D models, you can recognize their motions. |
| Q5 | You are able to improve your motion by the advice. |
| Q6 | Advice is easy to understand. |
| Q7 | Putting data gloves on is easy to do. |
| Q8 | Data gloves are burdensome to perform. |
| Q9 | The system is easy to use. |
| Q10 | It takes time to remember how to operate the system. |
| Q11 | The system is suitable to learn fingerspelling. |
| Q12 | You can continue learning through the system. |
| Q13 | You want to use the system again. |
| Q14 | Could you describe any functions which are needed for the system? |

### *6.1. Results of 5-point scale evaluation questionnaire*

The subjects, eight 20's students, answered the questionnaire. Table 3 shows averages and variance of each question.

Table 3. Results of the questionnaire

| No. | Average | Variance |
|---|---|---|
| Q1 | 3.1 | 0.61 |
| Q2 | 3.6 | 0.48 |
| Q3 | 3.0 | 1.0 |
| Q4 | 3.9 | 1.4 |
| Q5 | 3.8 | 1.7 |
| Q6 | 4.0 | 1.0 |
| Q7 | 4.8 | 0.19 |
| Q8 | 4.1 | 0.86 |
| Q9 | 4.4 | 0.48 |
| Q10 | 4.6 | 0.23 |
| Q11 | 3.8 | 1.7 |
| Q12 | 3.6 | 1.5 |

| | | |
|---|---|---|
| Q13 | 3.75 | 0.93 |
| Q14 | 3.1 | 0.61 |

*6.2. Result of free description*

In Q14, we asked subjects about functions they needed. As a result, there were few opinions about developing new functions. Therefore, we assessed that the functions of the system satisfied learners to be supported for learning fingerspelling.

*6.3. Subject's comments*

Answers obtained in the blank for free-comments provided at the end of the questionnaire are as follows.

- I get flustered, because when the 3D models overlays, it is hard to recognize which motion is the correct one.
- I think that you should change the process of multiple viewing.
- I would like to compare my motions before and after learning.

**7. Consideration**

*7.1. About data gloves*

From the high score about using data gloves by the evaluation, we can say the problem from previous study was solved [1]. In another hand, there was an opinion that the system was not able to capture finger motion precisely by the data gloves. In this case, it seems that the system used the data from the second joint instead of first joint to reproduce the first joint motion, since the data gloves cannot measure the angles of first joints. Even the learner bent only the second joint, the first joint of the 3D model was also bent, so sometimes the system was not able to reproduce the hand shape as precisely as the learner expected. In addition, the size of learner's hand depended on the learner, so that data gloves didn't fit the learner with small hands. In the future, we need to investigate the influence of differences among learners.

*7.2. About 3D models*

As one of the results about 3D models, we need to reconsider the way to indicate models, and coordinate models.

From the low score about Q3, it seems that the way of reproduction was not suited for indicating error. We think it was a reason that overlaying skilled person's and learner's models made it difficult for the learner to recognize each model, so that the learner was not able to understand a motion precisely. We propose a solution that the system will present only learner's model, and will emphasize the wrong point as feedback.

About 3D models, it was possible to select a 3D model corresponding to the size of the hand of each learner so that they were able to recognize the shapes of a 3D model easily. Therefore, a function for selecting an appropriate one from the models with different sizes by calibration data is necessary.

*7.3. About the function of advice*

In the comments by free description, the score of motions made the learner motivated to learn the motion. On the other hand, advice by text was not useful for them, so that they make sure how they improve their motion. To improve the quality of advice, we will examine the expression of the advice. In addition, we will send out questionnaires to check what kind of advice will be needed for sign language learners.

### *7.4. About the function of diagnose*

A function, which is able to get more information, is needed to advise the learner in more details. The information what we can get from the system is only similarity degree between the learner and skilled person. Therefore the system cannot compare each angle of skilled person's joints with each angle of learner's joints. That is why kinds of advice were limited, and utility of the function was decreased. To get more information, we plan to introduce Hidden Markov Model (HMM) which is advanced determination method to our system.

### *7.5. About the system*

On usability of the system, it got high score from Q9 and Q10. The reason was that the operation performed by the user was simplified. In addition, the questions related to the fact that the system is fit to learn fingerspelling got a roughly high score. There were some low scores on which the system is not fit to, however, that may be able to be responded by making an improvement about the way to show 3D models.

## 8. Conclusion

From the questionnaire, we got tasks to improve the way to show and overlaying 3D models.

In the future, we plan to make the system available to learn sign language expressions which are commonly used. Then, 3 dimension position sensors will be introduced to the system, which allows the system to get data of arms motions.

With the improvement of the system, it is needed to reconsider criteria for the evaluation, because the most important element is not to mimic the same motion as skilled person, but to learn a motion which a listener is able to understand. The system indicates a score for a similarity of a motion. However, in terms of physique differences of each learner, the learner is not able to synchronize his/her motion with a skilled person's motion precisely. As the result, even if the learner performs a motion which a listener is enough to understand, the learner gets a low score because of large difference from skilled person. For an establishment as a useful system, we have to figure out new criteria to evaluate whether the listener is able to understand the learner's motion of sign language.

## References


[1] Ryo K, Nobuhiko M, Shin K. Development of a sign language learning support system using data gloves. Technical report of Institute of Image Information and Television Engineers 2004;**28**:19-25.

[2] Mayu K, Yuhichiro M. Development of sign language training machine using Kinect. Technical Reports on Information and Computer Science from Kochi 2013;**5**:6-1.

[3] Masahito Y. The measurement of finger motion by a motion capture system using a magnetic position and direction sensor and reproduction of the 3D model. Graduation thesis of faculty of systems engineering, Wakayama university 2009.